

APPLICATION-SPECIFIC METHOD AND APPARATUS FOR ASSESSING
SIMILARITY BETWEEN TWO DATA OBJECTS

ABSTRACT OF THE INVENTION

The similarity between two data objects of the same type (e.g., two resumes, two job descriptions, etc.) is determined using predictive modeling. A basic assumption is that training datasets are available containing compatibility measures between objects of the first type and data objects of a second type, but that training datasets measuring similarity between objects of the first type are not. A first predictive model is trained to assess compatibility between data objects of a first type and data objects of a second type. Then, in one scenario, pairs of objects of the first type are compared for similarity by running them through the first predictive model as if one object of the pair is an object of the first type and the other object of the pair is an object of the second type. Alternatively, for each object in a set of objects of the first type, the first predictive model is used to create a respective vector of compatibility scores against a fixed set of objects of the second type; these various vectors are then used to derive measures of similarity between pairs of objects of the first type, from which a second predictive model is trained, and the second predictive model is then used to assess the similarity of pairs of objects of the first type.